Neural correlates of auditory-somatosensory interaction in speech perception

Takayuki Ito, Vincent Gracco, David J Ostry

Neural correlates of auditory-somatosensory interaction in speech perception

Takayuki Ito[1,2], **Vincent L. Gracco**[3,2], **David J. Ostry**[3,2]

1 CNRS, Gipsa-lab, Univ. Grenoble-Alpes, 2 Haskins Laboratories, 3 McGill University
takayuki.ito@gipsa-lab.grenoble-inp.fr

Introduction

Speech perception is known to rely on both auditory and visual information. However, sound specific somatosensory input has been shown also to influence speech perceptual processing (Ito et al., 2009). In the present study we addressed further the relationship between somatosensory information and speech perceptual processing by addressing the hypothesis that the temporal relationship between orofacial movement and sound processing contributes to somatosensory-auditory interaction in speech perception. We examined **changes in event-related potentials in response to synchronous (simultaneous) and asynchronous (90 ms lag and lead) somatosensory and auditory stimulation compared to that observed for unisensory auditory and somatosensory stimulation alone.**

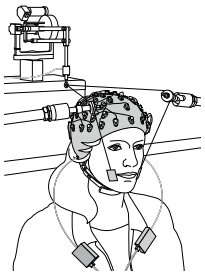
Method

Participants:

- 12 native speakers of American English (right-handed)

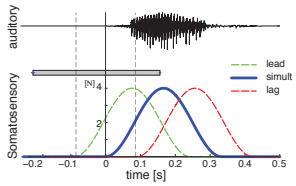
ERP recording:

- Event-related potentials recording from 64 scalp sites.
- Five stimulus conditions :
  Somatosensory stimulation alone (**SOMA**)
  Auditory stimulation alone (**AUD**)
  Three combined somatosensory and auditory stimuli (**90ms lead, Simult, 90ms lag**).
- 100 responses per condition.
- Stimulus presentation order was randamized.



Experimental setup

Stimulations:
Auditory: The synthesized sound between "head" and "had".
Somatosensory: Single cycle of 3 Hz sinusoidal skin stretch



Timing of auditory and somatosensory stimulations

Subject's task:
- Identify whether the speech sound was "head".
  68.5% were identified as "head".
- Gaze at a fixation point

Behavioral control:
Simultaneity judgments was examined by having participants identify when the two stimuli were perceived as occurring simultaneously.

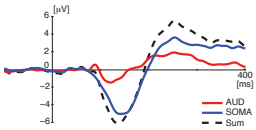
The average periods of **simultaneity** occurred **210.6 ± 3.1 ms lead** and **148.0 ± 3.9 ms lag** of somatosensory stimulation relative to the auditory stimulation.

Data analysis

Pre-preprocessing:
- Band-pass filtering (0.5-50 Hz)
- Re-referencing to the average across all electrodes.
- Bias levels adjustment relative to the pre-stimulus amplitude (-200 to -100 ms).
- Removal of exceptionally large responses (over ± 150 mV).

Somatosensory analysis:
(1) Align the data at **the somatosensory onset**

(2) Examine if: **Pair** = or ≠ **Sum** (SOMA + AUD)



Assumption: the summed ERP responses from the unisensory conditions (**Sum**) should be equivalent to the paired ERP (**Pair**), if neural responses to each of the unisensory stimuli are independent

Auditory analysis:
(1) Align the data at **the auditory onset**

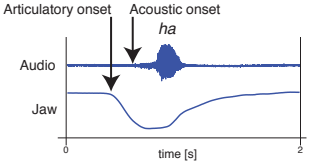
(2) Examine if **AUD** = or ≠ **AUD-like** (Pair - SOMA)

Assumption: "pair" ERPs with the removal of the somatosensory potentials (**AUD-like**) sould be equivalent to the auditory-alone ERP (**AUD**), if neural responses to each of the unisensory stimuli are independent

(3) Correlation analysis between N1 or P2 peaks and perceptual judgement of speech sound.

Discussion

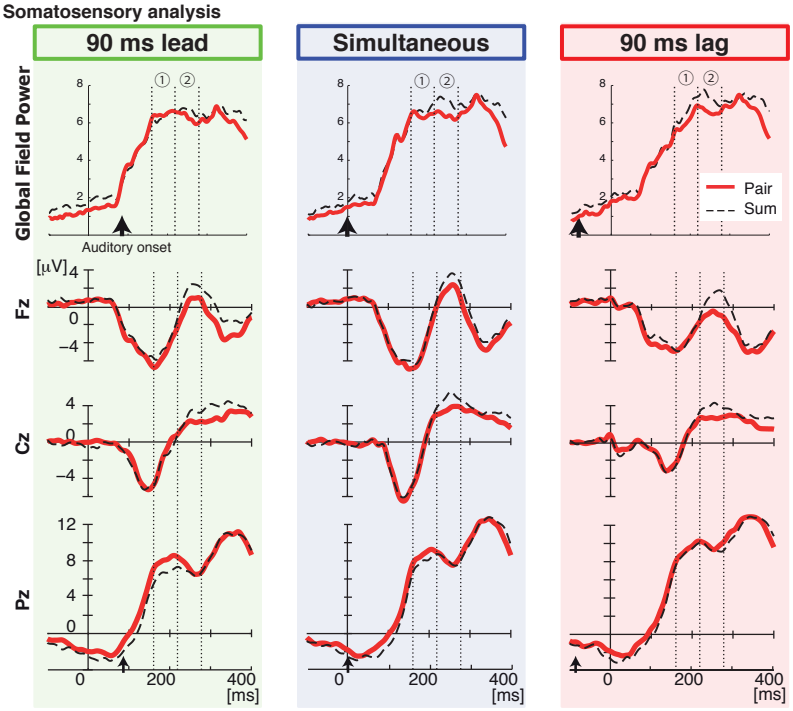
- A 200 ms of temporal asyncrony is considered to be the effective time-window for multisensory integration (Meredith et al. 1987, van Wassenhove et al. 2007). It is consistent with our results that dynamic modulation of the electrophysiological response was consistently induced at a temporal asyncrony within a 200 ms range.
- A dynamical modulation of multisensory processing at an electrocortical level was found at a range of 100 ms. The current finding suggests cortical processing is sensitive to temporal factors even within the time range at which events are behaviorally judged simultaneous.
- Given that articulatory motion generally precedes acoustic output of speech (Mooshammer et al. 2013), **the temporal relationship between orofacial somatosensory inputs and acoustic output in speech production might drive the neural process of somatosensory-auditory interaction.**

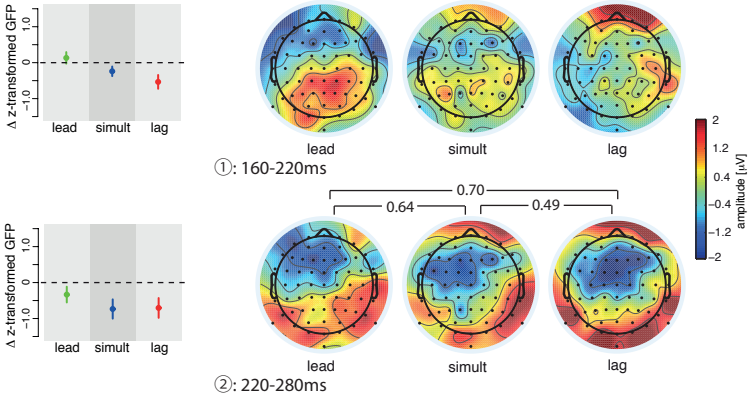


- Since multiple subcortical and cortical locations are involved in auditory–somatosensory interactions in non-speech processing (Stein and Meredith, 1993; Foxe et al., 2002; Lütkenhöner et al., 2002; Fu et al., 2003; Kayser et al., 2005; Murray et al., 2005; Schürmann et al., 2006; Shore and Zhou, 2006; Lakatos et al., 2007; Beauchamp et al., 2008), the present results reflect the contribution of different and distributed cortical sites in the somatosensory–auditory interaction in speech.

Results

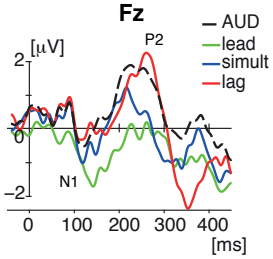
Somatosensory analysis

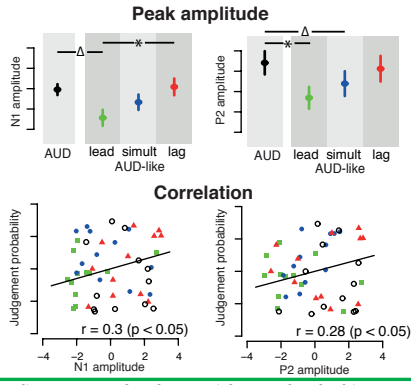




(1) Dynamical modulation according to stimulus timing in the periods of 160-220 ms
(2) Multisensory depression in the periods of 220-280 ms regardless of onset-time differences

Auditory analysis





(1) The modulation of peak amplitude of auditory event-related potensials was obserbed in 90 ms lead condition, but not in 90 ms lag condition.
(2) The peak amplitude correlated weakly, but reliably with the participant's judgement of speech sounds

Summary

(1) Cortical activity associated with somatosensory-auditory interaction was observed in event related potentials.
(2) Event related potentials associated with somatosensory-auditory interaction were modulated in response to changes in the relative timing of somatosensory and auditory stimulation.
(3) The results demonstrate a dynamic modulation of somatosensory-auditory convergence and suggest the contribution of somatosensory information for speech processing process is dependent on the specific temporal order of sensory inputs in speech production.

Acknowledgement
The research leading to these results has received funding from Notional Institute on Deafness and Other Communication Disorders, NIH, (R01DC012502 and R21DC013915), the Natural Sciences and Engineering Research Council of Canada and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 339152). We thank to Alexis Johns and Joshua Coppola for data collection and pre-processing.